DOCUMENT RESUME

ED 093 978 TM 003 834

AUTHOR Stetz, Frank P.
TITLE Toward Better Assessment of Student Achievement in Informal Educational Settings.
PUB DATE [Apr 74]
NOTE 17p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS *Academic Achievement; Achievement Rating; *Criterion Referenced Tests; *Evaluation; Literature Reviews; *Open Education; Student Records; *Testing

ABSTRACT

The research literature on open education has reported various studies describing and qualifying the term "open" in education and in attitudes of teachers involved in such programs. To date, very few large scale endeavors to assess student achievement in open education have been completed. Studies which have been done have not shown the hoped for increased gains over more traditional programs. This paper reviews the pertinent literature on these informal educational settings, proposes a more relevant assessment model for cognitive growth in such programs utilizing criterion referenced measurement, and proposes a more adequate system of reporting student achievement. (Author)

ED 093978
TM 003 834

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

# Toward Better Assessment of Student Achievement in Informal Educational Settings[1,2]

*Frank P. Stetz*
*University of Massachusetts*

In the past we have seen an abundance of innovative instructional models being implemented in our nation's schools. Most of these models have as one of their basic tenets the notion of individualized instruction. The rationale underlying these individualized models stresses the fact that children differ on such variables as interests, attitudes, intellectual development, environmental background, goals and so forth. More traditional instructional models have not typically taken into account these individual differences and perhaps this is why the schools are providing meaningful learning experiences for only a small portion of the children.

Some of the well-known individualized models include: Individually Prescribed Instruction (Glaser, 1968), Program for Learning in Accordance with Needs (Flanagan, 1967), Mastery Learning (Carroll, 1963, 1970); and what is most familiarly known in America as Open Education (Featherstone, 1968a, 1968b; Rathbone, 1971; and Barth, 1972).

While an abundance of literature is available on these new models, many problems remain. The testing component is particularly

---

[1] Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, April, 1974.

[2] The author would like to acknowledge the helpful comments and constructive criticisms of Ronald K. Hambleton on earlier drafts of this manuscript.

poorly handled in these new programs. Hambleton (1973) states:

> It is perhaps surprising to note...that the amount of information currently available on the testing methods and decision procedures for these programs is quite limited. It is this component that, in principle, facilitates the efficient movement of students through the instructional program [p. 3].

In particular, the assessment component in open (or informal) educational settings has been poorly defined. Barth (1969) states that "...the best way of evaluating the effect of the [open] school experience on the child is to observe him over a long period of time; the best measure of a child's work is his work." Although this is a logical approach toward assessment, in actual practice it would be difficult to "observe over a long period of time" a classroom of thirty children. Bussis and Chittenden (1970) imply that part of the reason for the absence of adequate assessment in open education is the lack of suitable measures on several of the student characteristics. In defense of better assessment, Walberg and Thomas (1972) believe that:

> Before...[open education] is expanded from the limited number of extant experimental settings in this country, administrators, teachers and parents quite properly should know if it leads to more learning, to higher levels of performance in reading...[etc.] [p. 207].

*Purposes*

A number of researchers (for example, Bussis and Chittenden, 1970) feel that a major reason for this poor assessment in open education has to do with the fact that the tests employed in the past have been used to order children according to more or less intelligence, more or less

readiness, and so on; that is, evaluators have used norm-referenced assessment.

Clearly required is a careful look at the testing and measurement needs of such informal educational models. As background to the study, there is a need to review the characteristics and reported research on these new programs. The purposes of this study are threefold: (1) to describe such models as open-space schools, open classroom schools, the integrated day approach, etc., helping to put further review and discussion of informal educational settings into the proper framework; (2) to review the pertinent literature on these open models concentrating on cognitive growth and assessment of children; and (3) to consider testing and measurement problems in open education, proposing a more attractive assessment model to measure and report cognitive growth utilizing criterion-referenced measurement.

*Descriptions of Selected Informal Educational Settings*

Brunetti, Cohen, Meyer and Molner (1972) define open-space schools to be:

> . . . composed of instructional areas without interior walls, ranging in size from two to over thirty equivalent classrooms. . . . Open-space schools . . . [can] consist of large open areas that can accommodate the entire student body and teaching staff [p. 86].

Brunetti, et al. go on to state that: "Teachers (in open-space schools) are no longer organizationally isolated but must cooperatively plan the activities of several groups of students. The task of planning becomes more complex, not only because of the number of students the team is

responsible for, but also because teams group and regroup students throughout the day and develop complex scheduling plans."

Open classroom schools are distinguished from open-space schools by their lack of vast amounts of architecturally open space. While open space is present to a limited degree in open classroom plans, schools of this nature do not require the integration of students and teachers characterized by open-space schools. Open classroom schoolrooms are usually self-contained and coordinated by one teacher with possibly the assistance of a teacher aid. These self-contained rooms serve as the home base in which students spend the majority of their time during the day. Featherstone (1971) states that open classrooms are flexibly arranged. They are divided into learning centers to provide for the simultaneous occurrence of several activities. Students are not limited to their seats to work, nor does the teacher remain in a fixed teaching area.

The integrated day or free day concept is best described by Weber (1971). She explains this approach by stating:

> In planning for the free (or integrated) day there is no separation of activities or skills and no separate scheduling of any one activity other than the fixed points . . . designed for all children in the school. As a result, one might see all aspects of the environment--reading, writing, numbers, painting, acting, music--in use at all times [p. 90].

From Weber's definition we can see that an integrated day approach may be the product of an open curriculum but does not necessarily have to be so. A traditional teacher may integrate her curriculum including arithmetic, language development, etc., without allowing pupils a choice in what will be the integrating factor.

British primary schools derive their name from the educational structure in England. At the present time, schooling is divided into primary and secondary schools. Primary schools encompass (although not always physically) both infant and junior schools. The usual age range of children attending these schools is five through seven for the infant school and eight through eleven for the juniors. Lady Bridget Plowden (Lady Bridget Plowden, et al., 1967) estimates that only one-third of the British primary schools can now be characterized as open. Consequently, to refer to open education and British primary schools synonymously is an error.

*Research on Cognitive Skills in Informal Education*

With regard to student achievement in open education, little substantial work has been reported. A few empirical studies have been made of the effects of architecturally open schools and experimental open classroom school programs on selected school outcomes.

Brunetti, et al. (1972) report that some studies have attempted to show that student growth in both affective and cognitive areas would be greater in open-space schools (Burnham, 1971; Kennedy & Say, 1971; Myers, 1971). Brunetti reports: ". . . no negative effects in either affective or cognitive growth have been shown to be associated with open space."

Pavan's (1973) review on research in the nongraded elementary school includes three studies examining the effects of student achievement in open-space versus traditional environments (Spencer, 1970;

Jeffreys, 1971; Warner, 1971). In all three cases no significant differences were found in student achievement between contrasting groups.

Gardner (1950, 1965, 1966) conducted longitudinal studies of the achievement of children in British integrated day classrooms. Evans (1971) concludes that Gardner's overall findings were favorable for the British integrated day classrooms compared to British traditional classrooms, although the traditional classrooms were not as carefully selected as the experimental, integrated day classrooms.

Tuckman, Cochran and Travers (1973) as part of their research on the effects of changing to open classroom schools compared the achievement of first through fifth graders in open and traditional schools using the California Achievement Test. Their results show that "standardized achievement was unaffected by the switch to open classroom; it was neither improved nor retarded."

*Assessment of Student Achievement in Informal Educational Settings*

We note from the review of cognitive growth research that typically the research has involved the use of standardized achievement tests. Results showing little or no significant differences between open and traditional classrooms were in the majority. The tests used in these studies were norm-referenced in nature and it has often been noted that the cognitive goals of open educational programs are not completely represented on standardized achievement tests. Also it should be noted that open education students are not frequently exposed to standardized achievement tests and hence their performance

may likely be hampered because of a lack of test sophistication.

A third argument against the use of norm-referenced tests in informal educational settings concerns its inadequacy as an individualized assessment tool. Open educators see norm-referenced testing as counterproductive to the goals of their programs. Their animosity stems not so much from an animosity to tests per se as from the fact that test results tend to turn the educator's attention away from individualized resources toward an attempt to categorize children (Bussis & Chittenden, 1970). While norm-referenced tests are of limited value for program assessment, they are even less useful for classroom monitoring. One alternative to improve program evaluation and classroom monitoring is provided by criterion-referenced testing. The assessment component of cognitive areas in open education could profit greatly if the proponents of such programs would look beyond inadequate testing strategies and integrate objective-based measurement in their required skill areas.

*A Proposal for Relevancy-Based Testing in Informal Education*

It is believed that open educators would display much less "animosity" toward testing and assessment if testing were more related to the specific decisions that teachers need to make; that is, if tests were constructed not to differentiate among children but to assess the actual state of affairs, to measure whether students have achieved the criteria by passing through the "threshold" from non-mastery of certain predetermined objectives to mastery of those objectives considered by

all to be important for development into thinking, intelligent adults.

What we are proposing is a criterion-referenced approach to the situation of assessing achievement in open education programs. Criterion-referenced tests have been defined in a variety of ways in the literature. (See, for example, Glaser & Nitko, 1971; Hambleton & Novick, 1973.) A very flexible definition has been proposed by Glaser and Nitko:

> A criterion-referenced test is one that is deliberately constructed so as to yield measurements that are directly interpretable in terms of specified performance standards. Performance standards are generally specified by defining a class or domain of tasks that should be performed by the individual. Representative samples of tasks from the domain are organized into a test. Measurements are taken and are used to make a statement about the performance of each individual relative to that domain [p. 653].

Hambleton, Stetz and Rios (1973) provide a decision-making framework for criterion-referenced measurement which would benefit teachers utilizing such tests. They state that testing is a decision-making process; that is, tests are given for the purpose of aiding in making decisions. "Decisions relating to mastery of instructional materials are best done with criterion-referenced tests." Test examinees in criterion-referenced testing situations consist of two mutually exclusive groups. One group is made up of examinees with high enough test scores to assume they have mastered the material; the second group is made up of examinees who did not achieve the minimum proficiency standard. The establishment of a cut-off score for determining mastery level is arbitrary and is primarily a value judgment.

This decision-theoretic approach toward testing is most

appropriate for the concerns facing open education teachers. To determine effectiveness of instruction and performance of individuals it is not necessary to rely upon fixed quota assessment strategies; most decisions made in open educational settings are quota free.

As outlined previously, criterion-referenced tests can be used to serve two purposes in open education. First, they can be used to evaluate the effectiveness of instruction. Norm-referenced tests given at the end of the school year or to compare instruction with some control group are usually inappropriate for making evaluative decisions on the effectiveness of instruction due to the fact that they are not designed to cover the instructional objectives. However, criterion-referenced tests are quite useful to the curriculum evaluator because of the specificity of the test results to the curriculum objectives.

Second, criterion-referenced tests can be used to provide very specific information on the performance levels of individuals on the instructional objectives. This information can be used, for example, to determine whether an individual has mastered particular objectives.

This new and more relevant approach to testing would provide more information to parents as well as teachers. Parents would be provided with performance-based data concerning what their children have accomplished in their open education learning experiences. Teachers would be provided with information necessary for the constant decision-making situations encountered in such settings.

To further clarify the intent of criterion-referenced measurement,

it should be noted that it would not be necessary to test all students at the same time concerning a particular objective or set of objectives. In fact, such a procedure would do much to destroy the essence of an open approach. Individual students or small groups working together on similar topics could be tested without interruption of classroom routine. Such criterion-referenced assessment questions could be integrated into the curriculum and included among the activities cards popular in most open education classrooms. The emphasis would be placed upon the assurance that what the children have covered is learned, and not the more traditional emphasis of testing with all its negative connotations.

The discussion so far has centered around the notion that our proposed use of instructional objectives and criterion-referenced test items measuring those objectives would be accepted by those in charge of informal educational programs. A point of fact is that such a proposal could generate a great deal of controversy with such proponents. The requirement of defining and stating objectives appears antithetical to such a movement. While a number of researchers (for example, Ebel, 1973) believe that it is inappropriate to invariably use instructional objectives in assessing achievement, it is possible to achieve a more realistic assessment of children's development in such areas where hierarchical structure and performance tasks are easily definable and desirable (i.e., mathematics and reading). This hypothesis should not be extended to more amorphous areas not relying upon a structure of hierarchical development nor to the integral

affective component of children's learning so prevalent in open education programs.

*Development of a More Relevant Reporting System*

Along with a better assessment of student achievement, a more representative and systematic approach to reporting student progress is needed. A traditional letter-grade approach to reporting student progress in such innovative programs is clearly outdated. Additionally, since most open education objectives are individualized, an approach which normalizes a class' scores into so many A's, B's, C's, etc., is completely out of place. Reporting systems utilizing performance objectives are basically more representative of student achievement, but still do not truly represent the individualized essence of most instruction in open education settings. Most performance-based reporting systems list a group of performance objectives that all students must master to reach criterion in a subject area. Columns are usually provided to check off and record the date when each objective is mastered. This assumes that <u>all</u> objectives in a particular subject are important for <u>all</u> students. This approach seems to lose the flavor of a truly open environment. In addition, very few systems such as the one just mentioned provide for credit to be given for those skills and mastered objectives that are completely unique to an individual learner. While it is believed that there are certain skills and objectives that should be mastered by all, most emphasis in open education relies upon individual differences in interests and

consequently mastery of unique skills and objectives.

A second argument favoring a more flexible and representative reporting system stems from the mood pervading much of American education today. The demand for accountability from both parents and administrators has had the effect of forcing teachers to account for their actions in the classroom. A reporting system which accurately depicts a profile of a child's accomplishments, whether they are required or elected, will help promote a clearer understanding of what children are learning in such informal settings.

What is being proposed is a more relevant reporting system for students in open education programs. This reporting system involves two main purposes: (1) to allow for adequate reporting of performance in those areas deemed important for all to master, and (2) to allow for credit to be given for those performances, objectives and tasks that are unique to an individual student. This system should accurately reflect what a student has learned from among the various alternatives available in open education programs.

This proposed reporting system could also be used to insure balance in a student's learning. Given that open education's assumptions rely upon a child being the prime planner of his learning experiences, it is necessary to monitor such activities. This system of reporting could act to balance the activities. With activities carefully monitored by this system, a teacher could easily detect areas in a student's program where gaps occur through lack of participation in certain required subjects. Therefore the proposed "checks and

balances" reporting system would provide checks to allow accurate bookkeeping of tasks accomplished, and balances in the curriculum to insure adequate coverage of material in various subject areas. Graphically, this proposed reporting system could possibly resemble a grid incorporating a system of check-offs for what a particular student has mastered.

*Expected Contribution to Education*

To date, most research studies dealing with open education have concerned themselves with describing and quantifying the term open education, and with teacher attitude and opinion toward such programs. What is being proposed is a rigorous study toward a body of knowledge concerning the students in these programs.

The time has come for something to be done on a large scale to evaluate objectively the effect open education has on the cognitive achievement of children in schools practicing such innovations. Little in the way of improvement and laudatory announcement can be made until a true assessment of the current state of affairs is made; that is, whether it "leads to more learning, to higher levels of performance in reading," etc.

This proposal on the use of criterion-referenced tests will hopefully point the way in the future toward the use of this new and more attractive approach to testing in open education. While previous undocumented attempts to study the question of student achievement in open education programs have come up with results "slightly" in favor

of a more traditional approach to education or no significant differences at all, it is believed that the wrong kinds of tests were used (those which purposely spread students out). If such procedures prove to be successful, a major advancement in the field of open education will be achieved.

References

Barth, R. S. Open education--assumptions about learning. Educational Philosophy and Theory, 1969, 1, 29-39.

Barth, R. S. Open education and the American school. New York: Agathon Press, 1972.

Brunetti, F. A., Cohen, E. G., Meyers, J. W., & Molner, S. R. F. Studies of team teaching in the open-space school. Interchange, 1972, 3(2-3), 85-101.

Burnham, B. Open education: Some research answers to basic questions. Orbit, 1971, 10(2), 22-26.

Bussis, A. M., & Chittenden, E. A. Analysis of an approach to open education. Princeton, N.J.: Educational Testing Service, 1970.

Carroll, J. B. A model of school learning. Teachers College Record, 1963, 64, 723-733.

Carroll, J. B. Problems of measurement related to the concept of learning for mastery. Educational Horizons, 1970, 48, 71-80.

Ebel, R. L. Evaluation and educational objectives. Journal of Educational Measurement, 1973, 10, 273-279.

Evans, J. T. Characteristics of open education: Results from a classroom observation rating scale and a teacher questionnaire. Final Report Project No. OEC-1-7-062805-3936, 1971.

Featherstone, J. The primary school revolution in Britain. The New Republic, August 10, September 2, and September 9, 1968. (a)

Featherstone, J. Experiments in learning. The New Republic, December 14, 1968. (b)

Featherstone, J. Schools where children learn. New York: Liveright, 1971.

Flanagan, J. C. Functional education for the seventies. Phi Delta Kappan, 1967, 49, 27-32.

Gardner, D. E. M. Long term results of infant school methods. London: Methuen & Company, Ltd., 1950.

Gardner, D. E. M. Experiment and tradition in primary schools. London: Methuen & Company, Ltd., 1966.

Gardner, D. E. M., & Cass, J. E. The role of the teacher in the infant and nursery school. London: Pergamon Press, 1965.

Glaser, R. Adapting the elementary school curriculum to individual performance. In Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton, N.J.: Educational Testing Service, 1968.

Glaser, R., & Nitko, A. J. Measurement in learning and instruction. In R. L. Thorndike (Ed.), Educational measurement. Washington, D.C.: American Council on Education, 1971.

Hambleton, R. K. A review of testing and decision-making procedures for selected individualized instructional programs. Review of Educational Research, 1974, in press.

Hambleton, R. K., & Novick, M. R. Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, 1973, 10, 159-170.

Hambleton, R. K., Stetz, F. P., & Rios, A. R. The development of objective-based programs in occupational education. Paper presented at the annual meeting of the Northeastern Educational Research Association, Ellenville, New York, 1973.

Jeffreys, J. S. An investigation of the effects of innovative educational practices on pupil-centeredness of observed behaviors and on learner outcome variables. Dissertation Abstracts, 1971, 31, 5766-A.

Kennedy, V. J., & Say, M. M. A comparison of the effects of open-area versus closed-area schools on the cognitive gains of students. Educators' Report and Fact Sheet, 1971, 8(4).

Myers, R. E. A comparison of the perceptions of elementary school children in open-area and self-contained classrooms in British Columbia. Journal of Educational Research and Development, 1971, 100-106.

Pavan, B. N. Good news: Research on the nongraded elementary school. The Elementary School Journal, 1973, 73, 333-342.

Plowden, L. B., et al. Children and their primary schools: A report of the Central Advisory Council for Education. London: Her Majesty's Stationery Office, 1967.

Rathbone, C. H. (Ed.) Open education: The informal classroom. New York: Citation Press, 1971.

Spencer, R. L. The development and exploratory application of an observational approach for studying the behavior and the behavior settings of individual students within elementary schools. Dissertation Abstracts, 1970, 31, 568-A.

Tuckman, B. W., Cochran, D. W., & Travers, E. J. Evaluating the open classroom. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, February 1973.

Walberg, H. J., & Thomas, S. C. Open education: An operational definition and validation in Great Britain and the United States. American Educational Research Journal, 1972, 9, 197-208.

Warner, J. B. A comparison of students' and teachers' performances in an open area facility and in self-contained classrooms. Dissertation Abstracts, 1971, 31, 3851-A.

Weber, L. The English infant school and informal education. Englewood Cliffs, N.J.: Prentice-Hall, 1971.